L3: Entry 27 of 97 File: USPT Aug 13, 2002

DOCUMENT-IDENTIFIER: US 6434561 B1
TITLE: Method and system for accessing electronic resources via machine-readable data on intelligent documents

Abstract Text (1):
A method of accessing electronic resources via machine readable data embedded on a document which comprises compressing input data with a transmitter adapted to save a first bandwidth using a compression method adapted to minimize utilization of bandwidth by the compressed input data while retaining substantially all information content of the input data and appending a compression flag to the compressed input data indicative of the compression method enabling a receiver to decompress the compressed input data. The compression step further comprises utilizing a compression dictionary adapted to map the elements and strings of the input data to minimized representations having redundancies deleted. The compression dictionary may be appended to the compressed input data (as cleartext or cyphertext) under circumstances where a bandwidth occupied by the appended compression dictionary is less than the bandwidth saved by the step of compressing the input data. The compression dictionary may also be selected by the receiver independently from the transmitter independently indexes, pointer registration, application restricted subsets or customized according to the input data content. Also the input data may be encrypted, and an encryption flag appended which is indicative of the encryption method enabling decryption via public or private key cryptosystems as well as utilizing various authentication techniques such as digital signatures to ensure that the document was created by a licensed user.

Brief Summary Text (2):
The present invention relates to a secure and efficient method and system for embedding machine-readable and executable data in a printed document and linking them to networked computer resources.

Brief Summary Text (4):
Electronic documents and printed media both enjoy advantages in certain situations, and may coexist and be used interchangeably depending on the application. The recent use of dense two-dimensional bar codes such as PDF417 has allowed the encoding of electronic data in a bar code symbol and the printing of such bar code symbol on a document. This technology is in its infancy and it is desired to implement secure, efficient methods of transferring data in electronic form on a printed document, referred to herein as an intelligent document. This enables the linking of electronic files with print media.

Brief Summary Text (5):
In addition to enabling the printing and distribution of electronic documents embedded in printed media, it is desired to enable a user to be given access to networked resources through such machine-readable symbologies. That is, in addition to providing a complete electronic file as an intelligent document, it is desired to be able to grant access to a targeted user to files found on an external resource, such as a computer network such as the Internet. That is, although a user is able in theory to enter a URL (uniform resource locator) into a browser program to obtain the Internet-based resource, such data entry on a keyboard is less than desirable.

Brief Summary Text (6):
However, human readable printed source addresses, and especially URL's, are

particularly difficult to manually enter in software programs, such as web browsers, due to their length and use of complex and unfamiliar symbols. If the characters in an address are not entered exactly, retrieval is prevented or, in a limited number of cases, a legal but incorrect source is accessed. This is especially true when URLs incorporate foreign languages and/or complex query instructions to on-line databases, as is increasingly frequent in many web sites. In addition, the inability to type or otherwise manually enter symbolic address information due to either a disability or lack of training complicates use of on-line information resources such as the Internet and World Wide Web for millions of individuals.

Brief Summary Text (7):
Thus it would be highly desirable to develop a method which automatically links particular sections of printed matter appearing on documents to on-line resources, whereby a user could, with a minimum of effort or experience, access on-line resources located at a variety of URL'S. This concept is not limited to on-line resources, but is equally applicable to accessing a variety of electronic resources within the user's immediate network as well.

Brief Summary Text (9):
Thus, a method which would eliminate typing and allow users to directly link printed addresses and query scripts to electronic information sources would be highly desirable.

Brief Summary Text (10):
In many instances the providers of on-line resources would find it necessary, for both commercial and security reasons, to restrict access to only those users who are authorized through a variety of licensing schemes. Many authorization techniques are in existence such as those disclosed by U.S. Pat. No. 5,388,158, hereby incorporated by reference, however, none have been incorporated in a method which provides access to on-line and local resources via printed matter.

Brief Summary Text (11):
Thus it would be advantageous to provide access authentication of potential users prior to granting access to on-line resources as well as local resources in order to guarantee that only authorized users could obtain access to restricted information and that the document was in fact created by a licensed source.

Brief Summary Text (14):
In accord with the present invention a method of accessing electronic resources via machine readable data embedded on a document is provided which comprises compressing input data with a transmitter adapted to save a first bandwidth using a compression method adapted to minimize utilization of bandwidth by the compressed input data while retaining substantially all information content of the input data and appending a compression flag to the compressed input data indicative of the compression method enabling a receiver to decompress the compressed input data. The compression step further comprises utilizing a compression dictionary adapted to map the elements and strings of the input data to minimized representations having redundancies deleted. The compression dictionary may be appended to the compressed input data (as cleartext or cyphertext) under circumstances where a bandwidth occupied by the appended compression dictionary is less than the bandwidth saved by the step of compressing the input data. The compression dictionary may also be selected by the receiver independently from the transmitter independently indexes, pointer registration, application restricted subsets or customized according to the input data content. Also the input data may be encrypted, and an encryption flag appended which is indicative of the encryption method enabling decryption via public or private key cryptosystems as well as utilizing various authentication techniques such as digital signatures to ensure that the document was created by a licensed user.

Drawing Description Text (2):
FIG. 1A is a block diagram which illustrates a method for embedding machine-readable data comprising electronic resources on an intelligent document of the present invention.

Drawing Description Text (3):

FIG. 1B is a block diagram which illustrates a method for recovering the electronic resources from the machine-readable data from the intelligent document of the present invention.

Detailed Description Text (2):
FIGS. 1A and 1B illustrate a block diagram of the method and system for encoding, rendering, distributing, recovering and accessing electronic resources via embedded machine-readable data on an intelligent document 46. The process of the present invention is initiated by a user, an automated source, or a primary application program 12 with input data 14 consisting of data files, executable programs, pointers to stored information or other digital data having information content stored on a computer system or transmitted on a digital data network.

Detailed Description Text (3):
A data classification step id first performed by the system. By properly classifying the data to be encoded, the recipient of the intelligent document 46 can scan the machine readable symbol rendered thereon and the application associated with the transferred data file will be automatically invoked. Thus the input data 14 is applied to a classification step 16 which sorts and tags the input data 14 according to the corresponding primary application program 12 which was either used to generate the input data 14 or is most closely associated with it. This can be done automatically with software by reading and analyzing file extensions of the input data 14 which are then mapped to common application programs 12 by convention. For instance, a file with a .doc extension would be mapped to a Microsoft Word.RTM. application. As a result, the recipient's computer will automatically execute Microsoft Word.RTM. upon processing the intelligent document 46. The classification step 16 could also be accomplished by analyzing the actual content of the input data 14 and identifying the specific or class of primary application programs 12 with which the input data 14 would most likely be associated with. Techniques which analyze the content or syntax of the input data 14 by searching for either specific deterministic markers such as the presence of `getchar` for C source code, or heuristic signatures such as the frequency of { }. Alternatively, the classification step 16 could be performed manually via operator designation upon initialization of the system as well as in real time during operation.

Detailed Description Text (4):
Once the associated primary application program 12 is identified, the input data 14 can be supplemented with prefixes, suffixes, labels or commands which are effective to communicate with subsequent secondary application 18 programs in an application coding step 20. For example, a URL=command could be prefixed to a string known to identify a web page (URL=http://www.neom.com) in order to signal software in the subsequent secondary application program 18 to invoke a web browser and link to that site. The application coding step 20 could also be used to provide pointer registration 22 to a local or distributed retrieval address of the file being referenced. This would provide a means for automatically creating and listing indexes for on-line retrieval of information through the use of machine-readable codes.

Detailed Description Text (6):
In addition or in the alternative to a URL, an index may be encoded within the intelligent document that may be sent by the user's browser program to a lookup table or index database located on a networked resource such as the Internet. A URL correlated to the index would be returned to the user's browser, and the browser would then use the URL to retrieve the resource from the appropriate server computer. This provides even greater flexibility since the URL may change in the future, and the content provider need only change the mapping function in the lookup table so that the same index is mapped to and returns a different URL. Thus, by including pointers such as indexes and/or URLs, great flexibility is provided by the intelligent document system where needed. Co-pending U.S. patent application Ser. Nos. 08/967,383 and 09/023,918 teach particular applications of indexes and resource addresses embedded in machine-readable symbols, and are incorporated by reference herein. FIG. 1C is illustrative of the overall system that retrieves a networked resource from a URL embedded in a machine-readable symbol embedded in an intelligent document, which may be implemented advantageously in the intelligent document system of the present invention as further described herein.

Detailed Description Text (7):
After the data to be encoded is determined, an optional compression step is implemented so that large data files can be encoded, even if the size of the file is too large to fit within the machine-readable symbology being employed. That is, during the compression step 24 the input data 14 is compressed to reduce the amount of bandwidth required to represent and convey the information without unreasonable distortion in the information content. This can be accomplished using compression methods which reduce the amount of redundant information in a transmission by optimally coding data elements or strings of data elements (i.e. tokens). In order to use these tokens to restore the original message during decoding, a compression dictionary can be transmitted with the input data 14, in which case a minimal acceptable compression value is defined as that point where the overhead of sending the compression dictionary with the data 14 is less than the bandwidth saved using compression. In the alternative to transmitting the compression dictionary inband, the compression dictionary may reside on an accessible data source (i.e., known and available to the recipient) and correctly mapped to the compressed data during decompression. Since the application programs associated with the input data 14 are known from the classification step 16 an appropriate compression method and associated compression dictionary can be defined and made available to both the compression 24 and decompression 60 steps based upon general message classification (e.g. letter frequency in English for plaintext English messages) or application restricted subsets (e.g. optimized compression for multi-token strings such as http:// for messages classified as html files).

Detailed Description Text (8):
FIG. 2 illustrates use of application restricted subsets of compression dictionaries 24A. During the compression step 24 a transmitting means has access to a number of subsets of compression dictionaries grouped according to application programs which are most closely associated with potential input data 14. Based upon the application associated with the input data 14 as determined in the classification step 16 of FIG. 1A a pointer or index is directed to a compression dictionary within a subset defined by the corresponding application program. The value of the pointer or index is transferred via in-band 26 or out-band 26A resources to a decompression step 60 on FIG. 1B which then uses the pointer or index to determine the appropriate method for the decompression step 60. This method of dictionary registration between the transmitting and receiving means enables the transfer of compressed input data without the explicit transmission of the compression dictionary utilized.

Detailed Description Text (10):
FIG. 3 illustrates an embodiment for application of customized compression dictionaries. The input data 14 is initially applied to a standard non-customized compression method 24B in order for a customization parameter/coefficient and control logic block 24C to obtain a predetermined sample of the content of the input data 14 in order to analyze it and determine the optimal compression algorithm, coefficients and parameters to apply according to methods well known in the art. After a sufficient sampling time has elapsed the customization parameter/coefficient and control logic block 24C may then direct the output 24E of the compression method to be derived from the customized compression method 24D by applying the appropriate select signal 24F to the multiplexor 24G. Due to the closed loop structure of this method dynamic variations in algorithm, coefficients, and parameters may continue to be provided to the customized compression method 24D during operation in order to maintain optimal performance of the overall compression step 24. Using the method described the compression dictionary may be transferred either in-band with the compressed data, via out-band resources, or not at all.

Detailed Description Text (11):
In addition to token compression, semantic-based variable coding compression may be utilized, whereby raw text information is analyzed and represented symbolically, transmitted in-band and then expanded at the receiving end using a set of common mapping conventions. An example of this technique would be a file where eye color is encoded as a simple numeric or bit pattern value (e.g. 1=blue or 00000001=blue). Another example would be the substitution of large numbers of boilerplate application parameters in an application data file with a symbolically coded value which indicates the application and the specific configuration in use in a native

file. This could then be expanded through substitution during the decompression step 60, thereby saving an appreciable amount of bandwidth. A compression flag 28 is appended to indicate which compression method was used and thus which method to be used during the decompression step 60.

Detailed Description Text (21):
User to host authentication schemes identify users to computer systems. The purpose of this type of authentication is to provide users with services for which they are authorized, and to deny access to services for which they are not. Those services might include an interactive login session, networked access to the host's file system or access to electronic resources as in the present invention.

Detailed Description Text (26):
An alternative scheme that prevents such attacks is a one-time password system. Unlike authentication mechanisms based on static passwords, those based on one-time passwords are not at all endangered by cleartext password entry. Three popular one-time password mechanisms are Bellcore's S/KEY, handheld authenticators, and smart cards.

Detailed Description Text (29):
Handheld authenticators, also called handheld password generators or tokens, are small hardware devices that generate one-time passwords. Use of handheld authenticators is based on the premise that each one is uniquely associated with exactly one user in the host's authentication database.

Detailed Description Text (33):
Smart cards operate much the same as handheld authenticators, however, they comprise more complex circuitry such as a central processing unit (CPU), a clock, program read only memory (ROM), RAM, and nonvolatile RAM or electrically erasable program read only memory (EEPROM) which are used to store and retain the key during power shutdown. The smart card permits the use of long keys without user intervention beyond entry of the PIN. The smart card communicates directly with the challenging entity via a reader and with the user via parallel, serial or PCMCIA interfaces.

Detailed Description Text (36):
Since the KDC is the only arbiter of authentication it must present a highly reliable and secure system since a breach of its security represents a problem to all principals which utilize its services. Also the KDC can become a bottleneck or result in total breakdown of communication links between any principal since all communications must pass through the KDC. Additional KDC's may be implemented, however, this presents further problems related to synchronization, maintenance and security. In addition, since all principals must place inherent trust in the KDC, this generally results in reducing the size of the environment. The Kerberos authentication system is an example of a trusted third party authentication method and is described in detail in Hughes, Internet Security Techniques pp. 91-125, (1995), which is hereby incorporated by reference. The authentication step 36 of the present invention could either be incorporated or operated in conjunction with any of the aforementioned methods of authentication (i.e., passwords, handheld authenticators, smart cards, and trusted third parties, or the equivalent method well known in the art).

Detailed Description Text (38):
One method of efficiently linking the printed document to outside resources is to encode the printed document using bar code symbology as described in detail in U.S. Pat. Nos. 5,243,655; 5,399,846; 5,504,322; and 5,471,533 which are hereby incorporated by reference. A bar code is typically a linear array of elements that are either printed directly on an object or on labels that are affixed to the object. Bar code elements typically comprise bars and spaces with bars of varying widths representing strings of binary ones and spaces of varying widths representing strings of binary zeros. Many bar codes are optically detectable and are read by devices such as scanning laser beams or handheld wands. Other bar codes are implemented in magnetic media. The readers and scanning systems electro-optically decode the symbol to multiple alpha-numerical characters that are intended to be descriptive of the article or some characteristic thereof. Such characters are typically represented in digital form as an input to a data processing system for a

variety of applications.

Detailed Description Text (39):
U.S. Pat. No. 5,388,158, which is hereby incorporated by reference, discloses a method which secures a printed document against tampering or alteration. This invention contemplates the document in question being scanned to produce a digital signal which is compressed, encrypted, and coded as a two dimensional bar code or some other appropriate form of coding, which is then incorporated onto a label and affixed to the document. In one embodiment the signal representing the image is encrypted using a pubic key encryption system and the key is downloaded from a central location. This key maybe changed from time to time in order to increase security. To facilitate authentication the corresponding decryption key is encrypted with another key and incorporated on the card. To validate the document the encoded signal is scanned from the label decoded, decrypted, expanded and displayed. The card may then be authenticated by comparing the displayed representation of the image with the document.

Detailed Description Text (48):
Upon completion of the decompression step 60, all information originally contained in the input data 14 is regenerated including data files, executable programs, macros, pointers and application coding as output data 62. The output data 62 is made available to an Application Programming Interface (API) 64 which invokes and feeds subsequent secondary application programs 18 which perform further program invocations and document display. For example, the command URL=http://www.neom.com would be interpreted by the subsequent secondary application program 18 which then invokes the designated web browser and links and executes the web page and Common Gateway Interface (CGI) script originally conveyed with the input data 14.

CLAIMS:

1. A method of accessing electronic resources via machine readable data on a document, comprising the steps of: compressing input data with a transmitting means adapted to save a first bandwidth using a compression method adapted to minimize utilization of bandwidth by said compressed input data while retaining substantially all information content of input data; and appending a compression flag to said compressed input data indicative of said compression method thereby enabling a receiving means to decompress said compressed input data, wherein said step of compressing input data further comprises utilizing a compression dictionary adapted to map said elements and strings of said input data to minimized representations of said elements and strings comprising redundant elements and strings deleted.

10. The method of claim 6, further comprising transferring said customized compression dictionary separately from said compressed input data through out-band resources.

15. The method of claim 1, wherein said compression dictionary is fetched from on-line resources.

16. The method of claim 1, wherein said compression dictionary is cached in resources local to both said receiving means and said transmitting means.

17. A method of accessing electronic resources via machine readable data on a document, comprising the steps of: compressing input data with a transmitting means adapted to save a first bandwidth using a compression method adapted to minimize utilization of bandwidth by said compressed input data while retaining substantially all information content of input data; and appending a compression flag to said compressed input data indicative of said compression method thereby enabling a receiving means to decompress said compressed input data, wherein said step of compressing said input data further comprises token frequency compression of said elements and strings, whereby frequency of repetitive elements and strings is enumerated rather than copying said repetitive elements and strings in order to represent repetition.

18. A method of accessing electronic resources via machine readable data on a document, comprising the steps of: compressing input data with a transmitting means

adapted to save a first bandwidth using a compression method adapted to minimize utilization of bandwidth by said compressed input data while retaining substantially all information content of input data; and appending a compression flag to said compressed input data indicative of said compression method thereby enabling a receiving means to decompress said compressed input data, further comprising the steps of: encrypting said input data using an encryption method; and appending an encryption flag indicative of said encryption method whereby said encrypted data may subsequently be decrypted, wherein said step of encrypting said input data further comprises a public-key cryptosystem, wherein said access authentication technique utilizes passwords to verify that said document was created by a licensed source, wherein said access authentication technique utilizes one-time passwords.

19. A method of creating a document capable of accessing electronic resources, comprising the steps of: encoding a static file in a machine readable code; encoding in the machine readable code a reference to an electronic resource in a computer network, wherein the electronic resource is capable of being modified without modification of the reference encoded in the code; and embedding said machine readable code in a machine readable symbol on a document, further comprising the steps of: compressing said input data using a compression method adapted to minimize utilization of bandwidth by said compressed input data while retaining substantially all information content of input data so as to save a first bandwidth; and appending a compression flag to said compressed input data indicative of said compression method thereby enabling a receiving means to decompress said compressed input data.

20. The method of claim 19, further comprising the steps of: utilizing a compression dictionary adapted to map elements and strings of said input data to minimized representations of said elements and strings whereby redundant elements and strings are deleted; appending the compression dictionary to said compressed input data; and transferring said compression dictionary with said compressed input data under circumstances where a second bandwidth occupied by said appended compression dictionary is less than said first bandwidth saved by said step of compressing said input data.

24. The method of claim 20, further comprising transferring said compression dictionary separately from said compressed input data through out-band resources.

27. A method of creating a document capable of accessing electronic resources, comprising the steps of: encoding a static file in a machine readable code; encoding in the machine readable code a reference to an electronic resource in a computer network, wherein the electronic resource is capable of being modified without modification of the reference encoded in the code; and embedding said machine readable code in a machine readable symbol on a document, further comprising the steps of: encrypting said input data using an encryption method; appending an encryption flag indicative of said encryption method whereby said encrypted data may subsequently be decrypted; and authenticating access to on-line resources via an access authentication technique adapted to ensure that said document was created by a licensed user, wherein said step of encrypting said input data further comprises a public-key cryptosystem, wherein said access authentication technique utilizes passwords to verify that said document was created by a licensed source, and wherein said access authentication technique utilizes one-time passwords.

28. A method of accessing electronic resources from a machine readable document, said method comprising the steps of: scanning a machine readable code from a machine readable symbol on said machine readable document; decoding a static file from said machine readable code; decoding a reference to an electronic resource in a computer network from the machine readable code, wherein the electronic resource is capable of being modified without modification of the reference encoded in the code; using the reference to an electronic resource to access an online electronic resource; and combining data obtained from said online electronic resource with the file decoded from said machine readable code, further comprising the steps of: reading a compression flag included with said input data, said compression flag indicative of a compression method used to save a first bandwidth by minimizing utilization of bandwidth by said compressed input data while retaining substantially all information content of input data; decompressing said input data according to a decompression method selected to match said compression method; and utilizing a

compression dictionary adapted to map elements and strings of said input data from minimized representations of said elements and strings whereby redundant elements and strings are deleted.

29. A method of accessing electronic resources from a machine readable document, said method comprising the steps of: scanning a machine readable code from a machine readable symbol on said machine readable document; decoding a static file from said machine readable code; decoding a reference to an electronic resource in a computer network from the machine readable code, wherein the electronic resource is capable of being modified without modification of the reference encoded in the code; using the reference to an electronic resource to access an online electronic resource; and combining data obtained from said online electronic resource with the file decoded from said machine readable code, further comprising the steps of: reading an encryption flag appended to said input data, said encryption flag indicative of an encryption method used to encrypt said input data; decrypting said input data according to a method selected to reverse said encryption; authenticating access to on-line resources via an access authentication technique adapted to ensure that said document was created by a licensed user, wherein said step of encrypting said input data further comprises a public-key cryptosystem, wherein said access authentication technique utilizes passwords to verify that said document was created by a licensed source, and wherein said access authentication technique utilizes one-time passwords.

30. A method of accessing electronic resources from an intelligent document, comprising the steps of: encoding a static file in a machine readable code; encoding a reference to an electronic resource in a computer network in the machine readable code, wherein the electronic resource is capable of being modified without modification of the reference encoded in the code; embedding said machine readable code in a machine readable symbol on a document; scanning said machine readable code from said machine readable symbol on said machine readable document; decoding the static file from the machine readable code; decoding the electronic resource reference from the machine readable code; using the electronic resource reference to access an online electronic resource; and combining data obtained from said online electronic resource with the static file decoded from said machine readable code, further comprising the steps of: compressing said input data using a compression method adapted to minimize utilization of bandwidth by said compressed input data while retaining substantially all information content of input data so as to save a first bandwidth; appending a compression flag to said compressed input data indicative of said compression method thereby enabling a receiving means to decompress said compressed input data; reading a compression flag included with said input data, said compression flag indicative of a compression method used to save a first bandwidth by minimizing utilization of bandwidth by said compressed input data while retaining substantially all information content of input data; decompressing said input data according to a decompression method selected to match said compression method; and utilizing a compression dictionary adapted to map elements and strings of said input data from minimized representations of said elements and strings whereby redundant elements and strings are deleted.

31. A method of accessing electronic resources from an intelligent document, comprising the steps of: encoding a static file in a machine readable code; encoding a reference to an electronic resource in a computer network in the machine readable code, wherein the electronic resource is capable of being modified without modification of the reference encoded in the code; embedding said machine readable code in a machine readable symbol on a document; scanning said machine readable code from said machine readable symbol on said machine readable document; decoding the static file from the machine readable code; decoding the electronic resource reference from the machine readable code; using the electronic resource reference to access an online electronic resource; and combining data obtained from said online electronic resource with the static file decoded from said machine readable code, further comprising the steps of: encrypting said input data using an encryption method; appending an encryption flag indicative of said encryption method whereby said encrypted data may subsequently be decrypted; authenticating access to on-line resources via an access authentication technique adapted to ensure that said document was created by a licensed user; reading an encryption flag appended to said input data, said encryption flag indicative of an encryption method used to encrypt

said input data; decrypting said input data according to a method selected to reverse said encryption; and authenticating access to on-line resources via an access authentication technique adapted to ensure that said document was created by a licensed user, wherein said steps of encrypting and decrypting said input data further comprise a public-key cryptosystem, wherein said access authentication technique utilizes passwords to verify that said document was created by a licensed source, and